

## EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S35	11	("6985898" "6665677" "7007174" "20010054153" "6742001" "6853997" "6839714" "6829606" "20040078364" "20040172442" "7020651" "7188107").pn. AND S34	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/12 16:12
S39	11	S37 AND S38	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/12 17:43
S38	51	(john NEAR2 ripley).in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/12 17:43
S36	23	("6985898" "6665677" "7007174" "20010054153" "6742001" "6853997" "6839714" "6829606" "20040078364" "20040172442" "7020651" "7188107").pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/12 17:43
S34	51	(john NEAR2 ripley).in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/12 17:43
S40	1	S37 NOT S39	US-PGPUB; USPAT	OR	OFF	2007/09/12 17:44
S41	10	S37 AND (weight\$4 normaliz\$7)	US-PGPUB; USPAT	OR	OFF	2007/09/12 17:48
S43	7	S37 AND algorithm\$1	US-PGPUB; USPAT	OR	OFF	2007/09/12 17:58
S44	13	S38 AND (weight\$4 normaliz\$7) AND document\$1	US-PGPUB; USPAT	OR	OFF	2007/09/12 18:03
S42	9	S37 AND (weight\$4 normaliz\$7) AND document\$1	US-PGPUB; USPAT	OR	OFF	2007/09/12 18:03
S46	14	S38 AND (error warning)	US-PGPUB; USPAT	OR	OFF	2007/09/12 18:04
S45	5	S44 NOT S42	US-PGPUB; USPAT	OR	OFF	2007/09/12 18:04
S47	0	"6738759".pn. AND (error warning)	US-PGPUB; USPAT	OR	OFF	2007/09/12 18:05
S50	2	(similarit\$4 AND sql AND document\$1).clm.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 13:32

## EAST Search History

S49	2	(similarit\$4 AND sql AND document\$1).clm.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 13:32
S48	2	(similarit\$4 AND (normaliz\$6 NEAR8 similarit\$4) AND sql AND document\$1).clm.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 13:32
S51	24	(similarit\$4 AND (normaliz\$6 NEAR8 similarit\$4) AND document\$1).clm.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:26
S53	17601	707/3-5,10.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:31
S52	2	"6738759".pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:31
S37	12	("6985898" "6665677" "7007174" "20010054153" "6742001" "6853997" "6839714" "6829606" "20040078364" "20040172442" "7020651" "7188107").pn.	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:32
S54	12	("6985898" "6665677" "7007174" "20010054153" "6742001" "6853997" "6839714" "6829606" "20040078364" "20040172442" "7020651" "7188107").pn.	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:39
S58	19	S56 AND ((seed) NEAR4 document\$1)	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:50
S57	97	S56 AND ((anchor\$1 source\$1) NEAR4 document\$1)	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:50
S55	584	(similarit\$4 NEAR4 document\$1) AND S53	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:53
S63	65	S60 AND (normaliz\$7 NEAR4 (similarit\$4 score\$1 weight\$1 rank\$1))	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:54
S62	116	S60 AND (normaliz\$7)	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:54
S61	163	S60 AND "707"/\$.ccls.	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:54
S60	187	(similarit\$4 NEAR4 (rank\$3 score\$1 weight\$1)) AND S59	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:54
S56	295	S55 AND (similarit\$4 NEAR4 (rank\$3 score\$1 weight\$1))	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:54

## EAST Search History

S68	67	S67 AND (normaliz\$7 NEAR4 (similarit\$4 score\$1 weight\$1 rank\$1))	US-PGPUB; USPAT; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:55
S67	195	S66 AND (similarit\$4 NEAR4 (rank\$3 score\$1 weight\$1))	US-PGPUB; USPAT; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:55
S66	471	(similarit\$4 NEAR4 document\$1) AND ((anchor\$1 source\$1 seed) NEAR4 document\$1)	US-PGPUB; USPAT; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:55
S65	15	S64 AND sql	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:55
S64	60	S63 AND "707"/\$.ccls.	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:55
S59	443	(similarit\$4 NEAR4 document\$1) AND ((anchor\$1 source\$1 seed) NEAR4 document\$1)	US-PGPUB; USPAT	OR	OFF	2007/09/18 14:55
S69	60	S68 AND "707"/\$.ccls.	US-PGPUB; USPAT; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 14:56
S70	4	("20040030716" "20050234976"). pn.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2007/09/18 15:50

[Web](#) [Images](#) [Video](#) [News](#) [Maps](#) [Gmail](#) [more ▼](#)[Sign in](#)

Google

similarity search document normalized

Search

[Advanced Search](#)  
[Preferences](#)

New! View and manage your web history

Web

Results 1 - 10 of about 711,000 for **similarity search document normalized**. (0.09 seconds)**Similarity search engine for use with relational databases ...**

The system of claim 20, wherein each result dataset includes at least one **normalized document similarity** score, at least one **search document** name, ...  
[www.patentgenius.com/patent/6829606.html](http://www.patentgenius.com/patent/6829606.html) - 144k - [Cached](#) - [Similar pages](#)

**Evaluating Strategies for Similarity Search on the Web**

In particular, given a query **document**, a **similarity-search** algorithm should ..... are the corresponding **normalized** frequencies (after scaling with the NMDF ...  
[www2002.org/CDROM/refereed/75/](http://www2002.org/CDROM/refereed/75/) - 90k - [Cached](#) - [Similar pages](#)

**CIPO - Canadian Patent Database - Claims - 2497191**

The system of claim 32, wherein the **normalized document similarity** scores ... one **normalized document similarity** score, at least one **search document** name, ...  
[patents1.ic.gc.ca/claims?patent\\_number=2497191&language=](http://patents1.ic.gc.ca/claims?patent_number=2497191&language=) - 27k - [Cached](#) - [Similar pages](#)

**CIPO - Canadian Patent Database - Claims - 2543397**

REMOTE SCORING AND AGGREGATING **SIMILARITY SEARCH** ENGINE FOR USE WITH ... more database management systems to determine **normalized document similarity** scores ...  
[patents1.ic.gc.ca/claims?patent\\_number=2543397&language=](http://patents1.ic.gc.ca/claims?patent_number=2543397&language=) - 23k - [Cached](#) - [Similar pages](#)

**Similarity search engine for use with relational databases ...**

computing a **normalized document similarity** score having a value of between 0.00 and 1.00 for each **search document** in each query command for indicating a ...  
[www.freepatentsonline.com/6829606.html](http://www.freepatentsonline.com/6829606.html) - 146k - [Cached](#) - [Similar pages](#)

**[PDF] A search strategy for large document bases**

File Format: PDF/Adobe Acrobat - [View as HTML](#)  
end **Search**. The above procedure computes the **document-query similarity** for .... in effect computes the **document** term weight as the intra-**document normalized** ...  
[cajun.cs.nott.ac.uk/compsci/epo/papers/volume1/issue2/epdxl012.pdf](http://cajun.cs.nott.ac.uk/compsci/epo/papers/volume1/issue2/epdxl012.pdf) - [Similar pages](#)

**Haveliwala, Taher;Gionis, Aristides;Klein, Dan;Indyk, Piotr ...**

**Similarity Search** on the Web: Evaluation and Scalability Considerations, ..... of the **document**, and  $f$  are the corresponding **normalized** frequencies (after ...  
[dbpubs.stanford.edu/pub/2001-8](http://dbpubs.stanford.edu/pub/2001-8) - 44k - [Cached](#) - [Similar pages](#)

**[PPT] Intelligent Information Retrieval and Web Search**

File Format: Microsoft Powerpoint - [View as HTML](#)  
Measuring **similarity** of two texts is a well-studied problem. ...  $f_{ij}$  = frequency of term  $i$  in **document**  $j$ . May want to **normalize** term frequency ( $tf$ ) by ...  
[www.cs.utexas.edu/~mooney/cs391L/slides/text.ppt](http://www.cs.utexas.edu/~mooney/cs391L/slides/text.ppt) - [Similar pages](#)

**[PDF] OCRed document**

**to-document similarity search**; on the other hand, when the ... as a vector of words together with **normalized** term frequencies. ...  
[ieeexplore.ieee.org/iel5/7762/21329/00989494.pdf?arnumber=989494](http://ieeexplore.ieee.org/iel5/7762/21329/00989494.pdf?arnumber=989494) - [Similar pages](#)



USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

 Search: ☒ The ACM Digital Library ☐ The Guide

+abstract:similarity +abstract:document +abstract:search

THE ACM DIGITAL LIBRARY


[Feedback](#) [Report a problem](#) [Satisfaction survey](#)
Terms used: **similarity document search**

Found 47 of 211,032

Sort results by

relevance

Display results

expanded form

☒ Save results to a Binder

☐ Search Tips

☐ Open results in a new window
Try an [Advanced Search](#)Try this search in [The ACM Guide](#)

Results 1 - 20 of 47

Result page: [1](#) [2](#) [3](#) [next](#)Relevance scale ☐ ☐ ☐ ☐ ☐

# 1 [Research track papers: Content-based document routing and index partitioning for scalable similarity-based searches in a large corpus](#)

Deepavali Bhagwat, Kave Eshghi, Pankaj Mehra

August 2007 **Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining KDD '07**

Publisher: ACM Press

Full text available: pdf(793.05 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

We present a document routing and index partitioning scheme for scalable similarity-based search of documents in a large corpus. We consider the case when similarity-based search is performed by finding documents that have features in common with the query document. While it is possible to store all the features of all the documents in one index, this suffers from obvious scalability problems. Our approach is to partition the feature index into multiple smaller partitions that can be hosted o ...

**Keywords:** distributed indexing, document routing, index partitioning, scalability, similarity-based search

# 2 [XML and information retrieval: a SIGIR 2000 workshop](#)

David Carmel, Yoelle Maarek, Aya Soffer

March 2001 **ACM SIGMOD Record**, Volume 30 Issue 1

Publisher: ACM Press

Full text available: pdf(401.67 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

XML - the eXtensible Markup Language has recently emerged as a new standard for data representation and exchange on the Internet. It is believed that it will become a universal format for data exchange on the Web and that in the near future we will find vast amounts of documents in XML format on the Web. As a result, it has become crucial to address the question of how large collections of XML documents can be sorted and retrieved efficiently and effectively. To date, most work on storing, indexi ...

# 3 [Chemoinformatics: an application domain for information retrieval techniques](#)

Peter Willett

July 2004 **Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '04**

Publisher: ACM Press

Full text available: pdf(134.13 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)


[Home](#) | [Login](#) | [Logout](#) | [Access Information](#) | [Alerts](#) | [Purchase History](#) | [Cart](#)

Welcome United States Patent and Trademark Office

☐ Search Results

BROWSE

SEARCH

IEEE XPLORE GUIDE

Results for "(((similarity &lt;near/4&gt; document &lt;near/4&gt; search))&lt;in&gt;metadata)"

☐ e-mail

Your search matched 6 of 1643271 documents.

A maximum of 100 results are displayed, 25 to a page, sorted by Relevance in Descending order.

## » Search Options

[View Session History](#)
[New Search](#)

## Modify Search


☐ Check to search only within this results set
Display Format: ☒ Citation ☐ Citation & Abstract

## » Key

IEEE JNL IEEE Journal or Magazine

IET JNL IET Journal or Magazine

IEEE CNF IEEE Conference Proceeding

IET CNF IET Conference Proceeding

IEEE STD IEEE Standard

 

- ☐ 1. **Applying the branch and bound technique to document similarity search**  
 Furuse, K.; Miura, T.; Ishikawa, M.; Chen, H.; Ohbo, N.;  
Communications, Computers and signal Processing, 2001. PACRIM. 2001 IEEE Conference on  
 Volume 1, 26-28 Aug. 2001 Page(s):331 - 336 vol.1  
 Digital Object Identifier 10.1109/PACRIM.2001.953590  
[AbstractPlus](#) | Full Text: [PDF](#)(472 KB) IEEE CNF  
[Rights and Permissions](#)
- ☐ 2. **On effective conceptual indexing and similarity search in text data**  
 Aggarwal, C.C.; Yu, P.S.;  
Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on  
 29 Nov.-2 Dec. 2001 Page(s):3 - 10  
 Digital Object Identifier 10.1109/ICDM.2001.989494  
[AbstractPlus](#) | Full Text: [PDF](#)(843 KB) IEEE CNF  
[Rights and Permissions](#)
- ☐ 3. **Finding themes in Medline documents-probabilistic similarity search**  
 Shatkey, H.; Wilber, W.J.;  
Advances in Digital Libraries, 2000. ADL 2000. Proceedings. IEEE  
 22-24 May 2000 Page(s):183 - 192  
 Digital Object Identifier 10.1109/ADL.2000.848381  
[AbstractPlus](#) | Full Text: [PDF](#)(284 KB) IEEE CNF  
[Rights and Permissions](#)
- ☐ 4. **Content-Based Clustered P2P Search Model Depending on Set Distance**  
 Jing Wang; Shoubao Yang;  
Web Intelligence and International Agent Technology Workshops, 2006. WI-IATW. 2006 IEEE/WIC/ACM International Conference on  
 Dec. 2006 Page(s):471 - 476  
 Digital Object Identifier 10.1109/WI-IATW.2006.53  
[AbstractPlus](#) | Full Text: [PDF](#)(298 KB) IEEE CNF  
[Rights and Permissions](#)
- ☐ 5. **Digital library access via image similarity search**  
 Heesch, D.; Pickering, M.J.; Howarth, P.; Yavlinsky, A.; Ruger, S.;  
Digital Libraries, 2004. Proceedings of the 2004 Joint ACM/IEEE Conference on  
 7-11 June 2004 Page(s):412